Improving Social Media Content Moderation: An Examination of the Flow of Online Conspiracy Theories across Varying Platform Models and Policies

By
Caroline Lanteri

Honors Thesis
Department of Public Policy
The University of North Carolina at Chapel Hill

April 27th, 2022

Advisor: Dr. Joaquin Alfredo-Angel Rubalcaba
Second Reader: Dr. Daniel Gitterman

Abstract

Misinformation, disinformation, and conspiracy theories have become increasingly prevalent on social media in recent years. This research focuses on two conspiracy theories that have resulted in deadly offline violence and their spread across social media platforms. It aims to uncover the ways in which social media platform models impact the evolution of conspiracy theories and identify areas for the improvement of content moderation. Through a case study approach, this thesis maps the evolution of the Great Replacement and Stop the Steal conspiracy theories. The Great Replacement refers to the idea that there is a deliberate plot to replace or rid of the world's white population. Stop the Steal claims that election fraud in the 2020 U.S. presidential election removed votes for former president and purported rightful winner, Donald Trump. The findings indicate that the massive reach of mainstream social media platforms and lack of oversight on fringe social media platforms create a flow of conspiracy theories from mainstream to fringe platforms. Extremist groups have a particular foothold on fringe platforms, and the conspiracy theories show a sharp escalation in extremism when they reach fringe platforms. Key actors on mainstream platforms, including right-wing news media outlets, draw in users to the conspiracy theory ideology and connect them to organized communities, where they are further directed toward fringe platform communities and greater radicalization. These findings provide insight related to opportunities for improved content moderation, including greater scrutiny of the most popular platform users, news media outlets, previously problematic groups, and especially linked content such as websites, blogs, or other platforms. Social media content shapes public discourse, and platforms have both the opportunity and the social responsibility to mediate the ways that this can be dangerous and even inspire offline violence.

Acknowledgments

I would like to thank my thesis advisor Dr. Joaquin Rubalcaba for his ongoing support and mentorship throughout this project, the many conversations that guided me toward this particular research question, and his guidance in structuring my case study approach. I would also like to thank my second reader Dr. Daniel Gitterman for his unwavering enthusiasm and encouragement throughout this difficult process. This thesis committee provided me with structure and feedback that allowed my idea and interest in social media culture to become a body of research. I am grateful to the UNC Department of Public Policy and Honors Carolina for providing me with academic communities that pushed me to be a better student, learner, and writer, and look forward to carrying these skills with me outside of Carolina.

Table of Contents

Chapter 1: Introduction **5**

Question *6*

The Cases: The Great Replacement and Stop the Steal *6*

Background and Context *7*

Policy Significance *9*

Chapter 2: Literature Review **10**

Conspiracy Theory Evolution Frameworks *10*

Conspiracy Theories on Social Media *12*

Policy Considerations *13*

Chapter 3: Methodology **17**

Selection of Cases *18*

Advantages *18*

Limitations *19*

Chapter 4: Findings **20**

Roots in News Stories *20*

Key Actors *22*

Repurposing of Old Communities *24*

Keywords and Coded Language *24*

Flow from Mainstream to Fringe *27*

Foreign Intervention *31*

Flowcharts *31*

Chapter 5: Conclusion **33**

The Impact of Network Models on the Evolution of Conspiracy Theories *33*

Policy Implications *33*

Suggestions for Future Research *35*

Chapter 1: Introduction

Social media platforms are the basis of communication in today's world, with a subsequent role in the dissemination of information. Their content shapes public discourse. Across the world, 4.48 billion people use some form of social media. That is 56.8%% of the world's total population, or 93.33% of global internet users (Dean, 2021). In the U.S., an estimated 72% of adults use at least one social media platform, and the majority visit the site at least once a week (Pew Research Center, 2021). Online misinformation and disinformation are growing threats because they can reach billions of people in an instant, and with real-world consequences.

Misinformation is defined as "incorrect or misleading information," whereas disinformation is "false information deliberately and often covertly spread in order to influence public opinion or obscure the truth" (Merriam Webster, 2021). This false content can take the form of incorrect data, manufactured news stories, and even conspiracy theories. It can be expressed in a number of forms with varying degrees of deception and harm: satire and parody, false connections, misleading content, false content, imposter content, manipulated content, and fabricated content (University of Michigan, 2021). The term "disinformation" is ten times as common in media headlines as it was five years ago, suggesting that the concern with false online content has shifted from a focus on mediating non-deliberate misinformation to an urgent addressal of organized disinformation campaigns (Nunberg, 2019).

Social media algorithms also contribute to the spread of misinformation, especially inflammatory content. They work to optimize viewership to amass profits by promoting extreme

content and creating echo chambers of information that cater to the preferences and affiliations of users.

Question

My thesis asks the following research question: How do varying platform models impact the evolution of conspiracy theories or disinformation campaigns across social media? This thesis examines how two different conspiracy theories or disinformation campaigns were deployed across social media platforms. These two instances are termed The Great Replacement and Stop the Steal and will be used as case studies. They will be examined across the social media platforms that they utilize the most, including both mainstream platforms such as Facebook, Instagram, YouTube, Twitter, Reddit, and fringe platforms, which in these cases either catered to or became overrun with extremist conservative users. My overarching hypothesis is that this content manifests differently on different platforms, particularly between mainstream and fringe platforms, but on each platform takes deliberate and tactical advantage of the network model and content moderation model.

The Cases: The Great Replacement and Stop the Steal

In 2019, an FBI intelligence bulletin first described "conspiracy theory-driven domestic extremists" as a growing threat. Advances in technology have heightened this risk as conspiracy theorists are now able to quickly reach millions of people via online platforms.

The "Great Replacement," first introduced by the French author Renaud Camus, is a white nationalist conspiracy theory that contends that a global elite is colluding to replace the white European population with non-European peoples, primarily through mass migration. It is a part of the white genocide conspiracy theory, which contends that there is a deliberate plot to rid of white people through mass migration, interracial marriage and miscegenation, and organized

violence (Bullens, 2021). This conspiracy theory is driven by ethnic hatred and has been employed time and time again to justify the white nationalist agenda, including terrorist attacks against Jewish people, Muslim people, Black people, and Hispanic people across Europe and the larger West.

The social media presence of the Great Replacement often takes the form of manifestos posted before terrorist attacks, which credit previous, related attacks with inspiration. Believers in this conspiracy theory also connect and communicate on social media platforms, sometimes in groups.

The Stop the Steal movement began with unfounded claims of election fraud in the 2020 U.S. presidential election and refers to coordinated efforts by supporters of former President Trump. This conspiracy fueled the attack on the U.S. Capitol on January 6th, 2021, which claimed five lives that day (Healy, 2021) and likely contributed to the later suicides of four police officers that had responded to the Capitol attack (Board, 2021). Disinformation can have actionable and dangerous consequences.

The rising popularity of far-right conspiracy theory movement QAnon, which played a large role in Stop the Steal, and other similar groups represents an American enthusiasm for baseless conspiracy theories that must not be left unchecked. It is a matter of national security, both domestically and internationally. In addition to domestic extremist groups, foreign countries are able to use this susceptibility as a tool to influence American politics and stoke division while promoting their own interests.

Background and Context

The largest technology companies operate with near-monopoly power and are often accused of employing anti-competitive practices. The four “big tech” companies: Apple,

Amazon, Facebook, and Google, have amassed their power by acquiring hundreds of companies over decades and now function as gatekeepers to the tech industry. As of September 2020, the combined valuation of these platforms is more than $5 trillion, which is more than a third of the value of the rest of the S&P 100 (Axios, 2020). They exhibit market dominance and an extreme concentration of the digital economy. Furthermore, A 16-month bipartisan investigation into competitive practices at these companies by the House Judiciary subcommittee on antitrust found that they operate with monopoly power that they have often abused. The investigation's report cites evidence that these companies "wield their dominance in ways that erode entrepreneurship, degrade Americans' privacy online, and undermine the vibrancy of the free and diverse press." This report recommends strengthening the Federal Trade Commission (FTC) and the Antitrust Division of the Department of Justice and promoting greater transparency and democratization of antitrust agencies (Subcommittee on Antitrust, 2020).

Members of Congress are working to pass antitrust legislation that will limit the power of these big tech companies. Six antitrust bills have recently been advanced by the Judiciary Committee. Some congresspeople contend that the breakup of companies like Facebook will improve the regulation of harmful online content.

Additional legislation of focus is Section 230 of the Communications Decency Act, which protects social media companies from liability for third-party content by classifying them as platforms rather than publishers. H.R. 2154, the Protecting Americans from Dangerous Algorithms Act, has been proposed in response to the algorithmic amplification of harmful online content. It would amend Section 230 of the Communications Decency Act to "hold large social media platforms accountable for their algorithmic amplification of harmful, radicalizing content that leads to offline violence" (Malinowski, 2021). The sponsors of the bill contend that

by opening the door to civil suits in cases where algorithmically-elevated social media posts lead to offline violence, H.R. 2154 would encourage social media companies to address the ways in which their platforms elevate these threats.

Policy Significance

These findings will help inform public policy and the platforms themselves on how to better moderate this type of content.

The digital world has become the underlying infrastructure for the exchange of communications and information. If research and policy intervention can be used to help platforms to better regulate dangerous, false content, it will change the public conversation. There will be a positive real-world impact in the realms of public safety, public health, mental health, and beyond. As new technology emerges, it is important to evaluate its social impact and enact policy to regulate its negative effects. It is clear that social media companies are not sufficiently regulated nor incentivized to fully mitigate misinformation and disinformation.

Chapter 2 will analyze the existing research on the success of online conspiracy theories or disinformation campaigns and the considerations for policy intervention. This literature review will situate my own research in the field of technology policy. Chapter 3 will describe methods and data and explain the choice of the following cases: The Great Replacement and Stop the Steal. Chapter 4 will analyze the results and findings of these case studies and Chapter 5 will provide policy recommendations to improve content moderation and make platform models more resistant to the spread of disinformation and conspiracy theories.

Chapter 2: Literature Review

This literature review will address three areas of research. It will first explore a number of frameworks for the evolution of conspiracy theories. Second, it will examine the current research about how social media platform models contribute to the spread of misinformation and conspiracy theories online. Third, it will provide an overview of policy considerations and suggestions for content moderation improvement and the rethinking of platform liability law. The chapter will end by discussing how this thesis will fill gaps in the existing research.

Conspiracy Theory Evolution Frameworks

The dual-pathway model of conspiracy belief describes the ideal situation for the emergence and success of a conspiracy theory as "separate but highly interconnected clusters of low levels of epistemic resources (comprehensive, accessible information) combined with high levels of epistemic need (a context of uncertainty) and low levels of socio-existential resources (trust) combined with high levels of socio-existential needs (a context of perceived threat)" (van Mulukom, 2021). Put more simply, it is when people face uncertainty without the solution of clear information, combined with feelings of perceived threat without the comfort of trust that someone will come to their aid- namely the government. The study describes humans as agents of information in social environments, seeking sources of authority to learn from. It notes dual-inheritance theories that suggest when humans gather information from their environment, they "crucially rely on their own senses and cognitive capacities (which are genetically inherited), and learn from trusted others (culturally inherited)" (van Mulukom, 2021). The study claims that authority is granted to different entities depending on the cultural context; on social media, mainstream authoritative institutes are often regarded as untrustworthy, and "authority is

transferred to those who are perceived to be credible, personalistic, and bring about a positive appraisal of oneself," playing into heuristics. (van Mulukom, 2021).

A cultural evolution theory approach explains that sometimes many conspiracy theories emerge about a particular event, indicating that certain events inspire conspiracy because of the nature of the event and not the nature of the evidence of conspiracy. Typically, this is related to uncertainty and fear surrounding the event and people searching for an explanation, an example being 9/11. This study's framework explains "conspiracy theories as the products of 'ordinary' cognitive and socio-cultural processes, while incorporating individual susceptibilities and wider contextual factors as influences on their cultural transmission and evolution" (Stubbersfield, 2021). This parts from more outdated explanations of conspiracism that describe those who believe in conspiracy theories as having a unique mindset and ascribe a psychiatric explanation rather than a cultural one. The cultural evolution theory approach, however, describes conspiracy theories as both a byproduct of socio-cultural processes and a product of individual minds. One particular pattern to note is the alteration of genuine news through social transmission to better fit cognitive biases (Stubbersfield, 2021).

Another writer describes conspiracy theory belief as an evolutionary trait. The article explains that "much of knowledge resistance is better understood as a manifestation of social rationality… Essentially, humans are social animals; fitting into a group is what's most important to us" (Klintman, 2019). This theory is compatible with others that describe the roots of conspiracy theories as situations of existential threat. This study also argues that the spread of the conspiracy theory is only successful when antagonistic outgroups are salient (van Prooijen, 2019).

A last theory of the evolution of conspiracy theories is the narrative theory, which again describes situations of low confidence and low trust as fueling the emergence and spread of conspiracy theories, calling this a threat narrative. It argues that the evolution of conspiracy theories relies on the "alignment of otherwise disparate domains of knowledge," for example news reports and social media, which are typically not harmonious (Shahsavari, 2020). It also notes that different conspiracy theories are often connected to one another and facilitate the connection of different groups and communities in a web of conspiracism (Shahsavari, 2020).

Conspiracy Theories on Social Media

Regarding social media specifically, Easterbrook (2021) argues in his dissertation that conspiracy rhetoric has molded itself to the social media landscape in ways that allow it to thrive. He examines three specific pieces of conspiracy content: Pizzagate images referencing a conspiracy theory about Hillary Clinton and a secret pedophile ring under a pizza shop, anti-Parkland or Parkland-denial headlines regarding the Parkland, Florida school shooting in 2018, and *Plandemic,* a conspiracy theory video about the Covid-19 pandemic. He argues that in each case, the rhetorical tools employed make specific use of the social media landscape. Conversely, conspiracy theory rhetoric shapes social media- "when it thrives, it makes social media a welcoming habitat for more conspiracy content to proliferate" (Easterbrook, 2021). As conspiracism becomes more familiar, it gains legitimacy and has now inserted itself into widespread political discourse in the United States. Easterbook proposes that the most promising intervention is to produce rhetorical responses to conspiracy content, though a labor-intensive effort, as well as integrate information literacy into school curriculums.

The Congressional Research Service (2021) published a report regarding social media misinformation and content moderation issues and the congressional role. The report states that

social media platform models can contribute to the spread of misinformation through two features: user networks and algorithmic filtering. The nature of social media algorithms is to maximize user engagement, which may amplify attention-grabbing misinformation. In 2018, an internal Facebook team presented slides to company executives stating that "[their] algorithms exploit the human brain's attraction to divisiveness," and warning that the algorithms would promote "more and more divisive content in an effort to gain user attention and increase time on the platform." Additionally, "Research using Twitter data from 2006-2017 has indicated that rumors or claims containing inaccurate information "diffuse significantly farther, faster, deeper, and more broadly" on social media than those containing accurate information" (CRS, 2021). People are naturally drawn to emotional and inflammatory content, and algorithms pick up on this tendency and fuel it further. This results in conspiracism becoming a successful part of the social media profit model.

Policy Considerations

Proposals for legislative reform in this area center around amending Section 230 of the Communications Decency Act to make social media platforms liable for user content. This has opened a debate over where to draw the line when it comes to preventing the spread of harmful misinformation without suppressing speech. Congress is considering the First Amendment implications of these legislative proposals. The report cites three factors that contribute to the amplification of misinformation on social media: the use of algorithms and data mining, the model of maximizing user engagement for greater advertising revenue, and the range of approaches to content moderation. Revenue from social media advertising makes up about 30% of all digital advertising revenue and amounted to roughly $36 million in 2019. Regarding content moderation practices, these rely on oversight by moderators, automated systems, and

users themselves. In 2020, Facebook paid a $52 million settlement to its content moderators, "who claimed to have experienced post-traumatic stress disorder from reviewing content on its sites" (CRS, 2021). There have also been many reports of Facebook's automated content moderation system removing harmless content and failing to remove certain illegal content.

In the past few years, warning labels have emerged as a potential mediator for mis/disinformation. Some platforms use them more heavily than others, and some, including Facebook and Instagram, use them instead of deleting posts flagged as misleading. There is not currently a consensus in the research regarding the efficacy of warning labels. In 2015, a paper warned of the "backfire effect," presenting evidence that social media users who saw fact checks were more likely to believe false claims. However, these findings have not been replicated, and other studies have shown decreased engagement with stories labeled as false. Another possible unintended consequence of widespread content warning labels is that users may interpret all unlabeled content as reliable, even when that is not the case. A study in 2020 by researchers from Harvard, Yale, MIT, and the University of Regina found that when some false content is labeled, users interpret all other content to be more accurate, calling this the "implied truth" effect (Quartz, 2020). Further studies suggest that fact check warnings may even cause people to double down on their beliefs (Marwick, 2018).

A more recent effort has been the promotion of reliable content. In 2020, Facebook, Twitter, TikTok, and Snapchat began promoting election information from reputable sources (Quartz, 2020).

In "Distributed intervention: networked content moderation in anonymous mobile spaces" by Nora A. Draper, her research into digital bystander intervention on the anonymous social media app YikYak finds a preference for moderation that relies on distributed

responsibility and collective action (2019). However, sometimes moderating misinformation may not be immediately clear to users, algorithms, or moderators themselves, and will have to be a matter of judgment (CRS, 2021).

In the Georgetown Law Technology Review, Tarleton Gillespie states that the tools of social media shape and center public discourse and culture. He argues that moderation itself is the essence of platforms, making them more than mere intermediaries or conduits of information, but rather facilitators with a responsibility to individuals and the public. The "intermediary" classification is what currently provides them protection under Section 230 of the CDA. Gillespie states that "The public problems we now face are old information challenges paired with the affordances of the platforms they exploit… platforms invoke and amplify particular forms of discourse, while moderating away others, all in the name of being impartial conduits of open participation" (2018). He argues for the introduction of legal parameters for content moderation, including the reconsideration of Section 230. He mentions that European legislators have gradually decreased the required time within which platforms must respond to online hate speech and terrorist propaganda. The European Commission implemented guidelines that require platform companies to be able to remove illegal content, particularly terrorist content, within one hour. Gillespie proposes that once a platform begins to curate or promote content, they become a hybrid between intermediary and publisher and should not be afforded all of Section 230's protections (Gillespie, 2018).

A report from the London School of Economics and Political Science's Truth, Trust, and Technology commission credits the "information crisis" in the UK to "five giant evils" among the public: confusion, cynicism, fragmentation, irresponsibility, and apathy (LSE, 2019). This report focuses on addressing problems of information reliability from social media, news media,

and politicians in concert with the sociological contributors to the information crisis. It proposes the establishment of an Independent Platform Agency with an observatory and advisory role. It also advocates for a statutory code on political advertising and considers heavier regulation in the long run, including potential platform liability for third-party content. For the larger public, it recommends greater media literacy education.

A broader proposal is that stricter anti-trust legislation or greater oversight by the Federal Trade Commission would force a fundamental restructuring of the largest social media companies (Seetharaman, 2020). A disruption of the internet's profit model could also change the way companies think about their algorithmic amplification of content.

Elon Musk has argued that Twitter's algorithm should be open source and give users the option of which algorithm to use or to use none at all. He has been historically critical of the company through tweets themselves, and in April 2022, he reached a deal to buy Twitter and strengthen its position on "free speech." The future of this deal and potential changes to the platform's guidelines remain unclear (MacMillan, 2022).

Many studies into online mis/disinformation and conspiracy theories examine how they spread and thrive in the context of social media as a whole. They examine algorithmic amplification, rhetorical tools, and other qualities that make this content successful online. There is a gap in the study of how this content finds ways to flourish on different platform models and within different contexts of content moderation. This information can help us better understand how to curb the spread of this harmful content.

Chapter 3: Methodology

A qualitative methodological approach will be used to analyze two instances of conspiracy theories or disinformation campaigns as critical instance case studies. The US General Accounting Office (GAO) refers to critical instance case studies as the most common application of case study methodology (1987). There will be just two selection sites, as critical instance case studies examine very few cases. This is because there is little interest in generalizability – analysis of the cases is important for better understanding the cases themselves, and making recommendations based on these particular findings. Under this approach, it is less important to generalize findings for other cases that are not being studied. Unlike GAO's typical application of critical instance case studies, these cases will not necessarily be used to determine cause and effect.

This research will employ thematic analysis under an evolutionary framework. This framework comes from the hypothesis that this content manifests very differently on mainstream platforms and fringe platforms, but in both situations takes deliberate and tactical advantage of the platform's network model and content moderation model. The platform's network model refers to the structure of the platform itself- the content users are able to post and view, the ways users are able to connect with others and join groups or view group pages, likes, comments, and other reactions, the algorithms that structure users' newsfeeds and searches, promoted content, and the structure of advertising. The content moderation model refers to the decisions that companies make about moderating content on their platforms, including removing content, warning labels, user ability to report content, standards for flagging content, and the entities they partner with to fact-check content.

This research will first attempt to flowchart the evolution of the two cases of conspiracy theories across social media to their culmination in offline violence. These flowcharts will help identify patterns in the way that conspiracy theories evolve, which will be elaborated on in detail in Chapter 4.

Selection of Cases

The two chosen instances of disinformation or conspiracy theories have both inspired some degree of offline activity, particularly violence, as these types of cases produce a large amount of available data. The Great Replacement and White Genocide have been credited as motivation for a number of terrorist attacks in the global West. In each instance, the white nationalist conspiracy theory is employed to direct blame at a particular minority racial or religious group, which becomes the intended target of an attack. While these attacks have typically been carried about by individual extremists, the theories are fueled by white nationalist academics who posit these theories as research. Stop the Steal also contributed to domestic terrorism when the U.S. Capitol was attacked on January 6, 2021. In this case, conspiracy group QAnon, as well as a large number of American politicians, fueled these claims of election fraud, many actually becoming involved in or leading the Stop the Steal movement.

The research will examine the presentation of these cases on the social media platforms that they most heavily utilized, including both mainstream and fringe social media platforms. This thesis will refer to Facebook, Instagram, YouTube, and Twitter as the major mainstream platforms, Reddit as a smaller mainstream platform, and the others as fringe platforms, which include 8chan, Parler, Telegram, MeWe, and Gab.

Advantages

The case study method provides an in-depth analysis of specific instances. The critical instance case study specifically "provides assurance that we have not prematurely overlooked important factors, that we have not been swayed by information from limited or perhaps biased sources, and that we have taken context into account, thus giving a fair and balanced picture of the situation" (GAO, 1987). The critical instance case study method provides for strong internal validity.

Limitations

The case study methodological approach does not provide external validity, or generalizability. As mentioned above, it is not a requirement to be able to generalize these findings. While the research question has a broader focus, the information from these specific critical cases can be used to suggest improvements on particular platform and content moderation models. These suggestions do not require an assurance of external validity. The specific applications of these cases do not need to be generalizable to other situations, they just need to highlight weaknesses and inform improvements in a way that may potentially have a positive impact on future situations. The cases can be thought of as observations for inductive reasoning.

Chapter 4: Findings

There are a number of patterns visible in terms of how these conspiracy theories evolve across mainstream and fringe platforms and how they make use of the interaction between these two types of platforms and each one's level of scrutiny in moderation. These findings will be explained in the following categories: roots in news stories, key actors, repurposing of old communities, keywords and coded language, flow from mainstream to fringe, and foreign intervention. Then flowcharts will be provided to further illustrate the evolution of the two cases, as well as a generalized version.

Roots in News Stories

An anonymous post on a pro-Trump Reddit forum called TheDonald first made the claim that Dominion voting machines had deleted millions of votes for Trump in the 2020 U.S. presidential election (Breland, 2020a). This idea was further promoted by the One America News Network, by Trump on Facebook and Twitter, and by the Trump Campaign's legal team. These claims were baseless and debunked by former U.S. cybersecurity chief Chris Krebs, who was then fired by Trump (Emerson, 2020a). There were more specific claims that an election software system company called Smartmatic had provided technology to Dominion and suppressed votes using technology provided by deceased former Venezuelan President Hugo Chávez, and that Smartmatic had been founded in Venezuela for the purpose of fixing elections. In truth, Smartmatic is a rival of Dominion, and although its founders are Venezuelan, it was founded in Florida and is now headquartered in England and no longer operates in Venezuela. An anonymous affidavit was submitted by someone in Texas claiming to have witnessed "the creation and operation of an electronic voting system in a conspiracy between a company known as Smartmatic and the leaders of conspiracy with the Venezuelan government" in a lawsuit

challenging the election results in Georgia. However, there are minimal details in the statement and no proof of any fraud in the 2020 U.S. election (Spencer, 2020). The increased number of mail-in ballots due to the Covid-19 pandemic was also adopted as an indication of voter fraud by right-wing conspiracy theorists, despite a lack of any proof.

Another conspiracy theory that went viral and was picked up by conservative media outlets on social media was the claim made by Trump lawyer Rudy Guiliani on Fox News and Twitter that deceased world champion boxer Joe Frazier was still voting in Pennsylvania. This possibly comes from the fact that voter rolls with registration information may contain incorrect birthdates or continue updating the age of a person who is deceased, but the Pennsylvania Office of Attorney General found no proof that any deceased person voted in the 2020 election. The deceased voter claims are a pattern in many elections in the U.S. but were particularly notable during the heightened tensions of the 2020 election (Lajka, 2020).

The White Genocide conspiracy theory is rooted in pseudoscience and ethnic hatred by academics that has been capitalized on by extremist right-wing politicians, hate groups, and individuals. The conspiracy theory has taken many different forms, directing blame for the supposed threat of white extinction towards Jewish people, Muslim people, Black people, working-class immigrants, Hispanic migrants to the U.S., African migrants to Europe, and even simply American Democrats, based on the situational reasoning. In one instance, Twitter users cited CDC racial breakdowns of fertility rates in the U.S., which showed declining white birth rates, as proof of a plot of white extinction by immigrant-born women. One tweet read, “mission accomplished for the social engineers” (Betuel, 2019). The CDC report was quickly spread on the alt-right blog InfoWars and used to promote and justify white nationalist ideology.

Another manipulation of news reports relates to land reform in South Africa. A number of killings of white farmers in South Africa have been equated to post-apartheid land redistribution policies, which are then referenced as evidence of a race war and "white genocide." Meetings and interviews between South African right-wing radicals, such as leaders of the Christian survivalist group the Suidlanders, and American right-wing news hosts and media personalities, eventually brought this conspiracy theory to then-President Trump, who tweeted about it and ordered then-Secretary of State Mike Pompeo to examine the issue (Pogue, 2019). The framing of South African land seizures as a potential focus of U.S. foreign policy gave credence to the conspiracy theory. Believers in the white genocide conspiracy theory also focus on advocating for restrictive immigration policy and against miscegenation. A HuffPost analysis found that Trump's Facebook page ran approximately 2,200 ads in 2019 "stoking xenophobia and warning of an "invasion" by immigrants — a revelation receiving attention in the wake of the mass shooting in El Paso, Texas," which was one of the attacks fueled by the white genocide conspiracy theory (Russo, 2019).

Right-wing news media is unique in that it often aligns with some of the extreme right-wing content on social media. This is an example of two narrative domains that align in a way that is atypical for news media and social media. The narrative theory of the evolution of conspiracy theories describes this alignment as important for conspiracy theories to thrive.

Key Actors

Influential actors emerge on social media typically because they represent a source of either authority or insider knowledge. In some cases, even anonymous postings that claim to be from sources of insider government or conspiracy theory knowledge are successful in garnering massive followings and support.

The Stop the Steal movement was fueled by prominent conservative politicians and right-wing news media correspondents who used Twitter as their primary social media. After Trump declared a premature victory and called for vote counting to stop, prominent right-wing social media users including "Republican Rep. Marjorie Taylor Greene, American Conservative Union Chairman Matt Schlapp, and far-right media personalities Dice and Michelle Malkin" began promoting #StopTheSteal (DFRLab, 2021). An analysis found that tweets using this hashtag between 9 p.m. ET Nov. 2 through 5 a.m. ET Nov. 4 2020 were most promoted by accounts used by "the Philadelphia GOP, Human Events publisher Will Chamberlain, Trump-supporting actor James Woods, Islamophobic activist Pamela Geller, conservative writer Carmine Sabia, and online influencers Posobiec and Mark Dice" (DFRLab, 2021).

Prominent figures such as former President Trump and attorney Sidney Powell also amplified baseless conspiracy theories from right-wing news sources such as Breitbart on Twitter and in press conferences that were later circulated across multiple social media platforms. There was heavy interaction between government accounts and those of media personalities. For example, the Arizona Republican Party Twitter account retweeted right-wing conspiracy theorist and media personality Ali Alexander's tweet, "I am willing to give my life for this fight" (DFRLab, 2021).

These findings echo previous studies of the evolution of conspiracy theories that credit uncertainty, feelings of threat, and loss of power as motivators for belief in conspiracy theories.

For the white genocide conspiracy theory, the key actors on social media seem to be right-wing news correspondents and media personalities who interact with members of extremist right-wing groups. These mainly include conservative media pundit Ann Coulter, Fox News host Tucker Carlson and his interviews with people like Afrikaners activist and AfriForum deputy

head Ernst Roest, and social media personality Mike Cernovich and his interviews with people like Suidlanders leader Simon Roche (Pogue, 2019). Conservative politicians with connections to these right-wing media personalities further popularize their content.

Repurposing of Old Communities

For Stop the Steal, a Facebook group called "Boot Pelosi" was repurposed and renamed "March for Trump" by members of Women for America First. The repurposing of old groups for new causes is a common tactic when the member demographic is likely to share similar support for a new cause. Anti-Covid-19 lockdown groups and previous campaign groups were also rebranded to support Stop the Steal and organize members to rally. A private, extreme Facebook group called "Hardcore Trump Nation" with 12,000 members was changed to read "Stop the Steal" two days after the election (Emerson, 2020a). Members of many Facebook groups were also encouraged to join "recidivist" groups in case the originals were removed by Facebook (Bond & Allyn, 2021). Social media allows those with similar mentalities to connect to one another over a specific topic and subsequently be guided to similar, and often more extreme ideas.

Keywords and Coded Language

Establishing terminology is important for the deliberate spread of conspiracy theories because it makes them more accessible to users that may be searching for them on social media, and it also makes use of social media algorithms that promote popular content and suggest related content to users.

For the Stop the Steal movement, the hashtag #StopTheSteal became a way for pro-Trump believers in the election fraud conspiracy theory to connect with one another. Pro-Trump Facebook groups such as "President Donald Trump Fan Club" and "Donald Trump is My

President" reposted this hashtag, reaching millions more people. Fringe news media outlets such as pro-Trump blog *The Gateway Pundit* began to use the slogan in articles, which were further circulated on social media. Some media personalities, such as Ali Alexander, even changed their Twitter names to include "#StopTheSteal" to make their affiliations immediately visible and attract more attention and followers. Hate groups such as the Proud Boys use specific terminology and phrases that help potential supporters or members find them on social media. For example, they refer to themselves as "Western chauvinists" and employ mottos such as "West is best," "Glorify the entrepreneur," "Venerate the housewife" and "I won't apologize for creating the modern world" (The Associated Press, 2018). On Parler, one of the major fringe platforms used for this misinformation campaign, promoted hashtags included #STOPTHESTEAL, #MAILINBALLOTS, and #HUNTERBIDENLAPTOP, all related to connected conspiracy theories regarding the election (Emerson, 2020b).

In 2018, #Whitegenocide was the most popular hashtag used by white nationalists on Twitter (Coaston, 2018). This rhetoric originated in the "White Genocide Manifesto" by David Lane, the convicted murderer of a Jewish radio host in 1984. Other phrases used to refer to similar theories include "the Kalergi plan," "Eurabia", "diversity is code for white genocide," "anti-racism is code for white genocide," "white extinction," "white replacement," "the great replacement," "immigrant invasion," and references to the "homeland" in relation to immigrants. Some of these are also identified by the academics or authors of specific theories, like Bat Ye'or for "Eurabia," which is the conspiracy theory that French and Arab powers are seeking to Islamise and Arabise Europe- a subset of the white genocide conspiracy theory.

Furthermore, "You will not replace us" and "Jews will not replace us" became rallying cries for white supremacists that transferred from social media to chants at the "Unite the Right"

rally in Charlottesville, Virginia, where a woman was killed by a neo-Nazi who drove his vehicle into a crowd of counter-protesters.

White genocide and related conspiracy theories have fueled many deadly attacks across the West. These are most often carried out by individual actors who post online or on social media preceding the attack. While some of the perpetrators' social media activity can be identified as extremist on their mainstream social media accounts, their manifestos are typically posted on fringe social media platforms such as Gab or 8chan or on their personal websites or blogs. These manifestos have also shown a pattern of building on or utilizing one another.

In 2011, Anders Behring Breivik killed 77 people in Norway in a mass shooting and bombing. On the day of the attacks, he posted a manifesto blaming Islam and feminism for the "cultural suicide" of Europe. He actually copied entire sections of the Unabomber Manifesto, replacing words such as "cultural Marxists" for "leftists" and "Muslims" for "black people" (van Gerven Oei, 2011).

In 2019, Patrick Crusius opened fire in a Walmart in El Paso, Texas, killing twenty people. Minutes before the shooting, he posted a manifesto to fringe platform 8chan, which praised the perpetrator of the Christchurch, New Zealand gunman, who had posted his own manifesto on 8chan before killing 50 people (Arango, 2019). It also followed a similar pattern to the Poway, California shooting, whose perpetrator also posted a manifesto on 8chan. Crusius' manifesto expressed anti-immigrant sentiment, where the Christchurch shooting targeted Muslim people at a mosque and the Poway shooting targeted Jewish people at a synagogue. All three manifestos employed similar white nationalist talking points, and all three attacks were carried out in 2019.

Whereas white supremacist academics published their content in newspapers and books and justified their claims as science or research, many white supremacists online justify their claims as political opinion, which they argue should always be protected speech.

The use of key terms, sometimes in the form of coded language, facilitates the connection of like-minded people on social media. This is amplified by platforms' algorithms, which navigate users toward content they are most likely to interact with. The social media profit model thrives off of views and interactions with content.

Flow from Mainstream to Fringe

Some misinformation and key spreaders have been removed from mainstream social media platforms, particularly Twitter. For example, Donald Trump's account @realDonaldTrump was banned from Twitter on January 8th, 2021, "due to the risk of further incitement of violence" (Twitter, 2021). In 2019, Twitter specifically outlined principles and approaches for the moderation of the accounts of world leaders. Generally, this states that world leaders' accounts will be monitored closely and tweets that violate the Twitter Rules will be removed unless they are of specific public interest value, in which case they will be placed behind a notice (Twitter, 2019). Twitter also de-platformed *The Gateway Pundit* in February 2021 and continues to ban news outlets and services that they determine distort stories and "violate the Twitter Rules on platform manipulation and spam" (Waterson, 2022). Trump was also banned from Facebook until at least January 7, 2023, following the January 6th insurrection, due to a "serious risk to public safety" and praising of violence (Bond, 2021). Similar to Twitter, Facebook employed an exemption to its speech policy for "newsworthiness," applying by default to posts by politicians. However, due to a recommendation by the external Facebook Oversight Board, Facebook decided to use greater scrutiny when weighing newsworthiness against the

potential risk of harm and disclosing these decisions. Another policy that exempts politicians' posts and paid ads from Facebook's third-party fact-checking was left in place (Bond, 2021).

Despite posts and users being removed, even a short amount of time that conspiracy theory posts or articles are accessible on mainstream social media can allow users to connect and move to fringe social media platforms, where they have the freedom to radicalize and mobilize with very little oversight. When Facebook banned the first Stop the Steal group on November 5th after the November 3rd election, it had already gathered 360,000 members. Before it was removed, a member wrote "I'm moving to Parler and suggest you all do too." A new user on Parler wrote, "Try to get as many as you can to join us here… We are using Facebook as a platform to bring the crowd." Many more users began to share the sentiment, and a new Parler account sharing misinformation called "Stop the Steal" quickly amassed over 100,000 followers. Conservative pundits like radio host Mark Levin even encouraged people to flock to Parler to avoid the "Big Tech censorship" of Facebook (Emerson, 2020b).

After Donald Trump called on the Proud Boys to "stand back and stand by" in a Fox News interview and these words were further circulated by reporters and commentators on social media, hundreds, perhaps thousands of users joined Proud Boys groups on fringe social media platforms. The white nationalist group was banned from Facebook, Instagram, and Twitter in 2018. They had previously been recruiting new members through private Facebook chat rooms (The Associated Press, 2018).

Fringe platforms are heavily used by militia groups as their lack of oversight lends the ability to organize violence. The OathKeepers private chatrooms and MeWe group chats, such as one titled "AMERICAN CIVIL WAR 2.0," discussed the potential for organizing violence after the election. In late November, an OathKeepers chatroom called for law enforcement officers

with firearm permits to attend election protests. In MeWe chatroom "American Patriots III%," the Three Percenter militia group also organized protest attendance. A Proud Boys chapter used Telegram to share violent imagery and promote the protests. Facebook groups such as "Texas for Trump" then further promoted these rallies by citing the attendance of prominent members of the aforementioned militia groups (DFRLab, 2021). These fringe platforms have vague guidelines prohibiting threats of violence but did not seem to be effective in removing any of this content. While these platforms call themselves "free speech havens," it is important to remember that incitement to violence is not protected speech in the U.S. (Emerson, 2020b).

A Facebook group called "STOP THE STEAL," founded by pro-Trump group Women for America First, gained hundreds of thousands of members in one day before it was shut down by Facebook and promoted multiple rallies across the country in an events section. The same day, Ali Alexander's StopTheSteal.us website received significant traction on Facebook, and the website was updated to prompt visitors to input their contact information to facilitate organization as well as include listings for protests across the country. While Facebook took greater action in the days following the election, some groups had been left unmoderated after being flagged internally as high risk for election misinformation months earlier, and further internal research had found that "Facebook's groups recommendations could quickly send users down partisan political rabbit holes" and that recommendations "devolved toward polarizing content" (Bond & Allyn, 2021). Closer to the election, Facebook rolled out "break the glass" measures where they designated the U.S. a high-risk location, reduced the visibility of posts deemed likely to incite violence, and restricted the growth of some political groups (Bond & Allyn, 2021).

The Reddit forum TheDonald, which hosted the first claims of Dominion voting machine fraud, was not banned for years despite repeated content that violated Reddit's speech policies on racism, bigotry, and disinformation. By the time it was removed, users had employed the forum to encourage traffic to a homemade forum that mimicked Reddit called TheDonald.win as well as to servers hosted by Discord to circumvent Reddit's ban (Breland, 2020b).

On December 12, 2020, there was a large pro-Trump election protest, where key actors took to mainstream platforms to praise and cover the events. Several people were stabbed and dozens arrested.

The organization for the January 6th, 2021 protest was even more extreme. Trump tweeted about a report alleging election fraud and added, "Big protest in D.C. on January 6" (DFRLab, 2021). Twitter added a warning label to this tweet, which read "This claim about election fraud is disputed" (DFRLab, 2021). However, other key actors replied to this tweet with website links such as TrumpMarch.com and StopTheSteal.us where these claims were reaffirmed to viewers. An administrator of 8chan shared a page with details for organizing travel to D.C. for protests on January 6th. A graphic with instructions for joining vehicle caravans heading to D.C. was circulated on a number of fringe platform communities. A Facebook group called "Red State Secession" sent its nearly 8,000 members to a website with information on travel routes to D.C. (Lytvynenko & Hensley-Clancy, 2021). Ali Alexander posted a YouTube video, which has now been deleted, encouraging people to register for protests and organize transportation on WildProtest.com (DFRLab, 2021).

While Trump and other conservative politicians continued to get the word out about the January 6 protest, militia groups continued to organize on fringe platforms. One post on Parler, which received hundreds of thousands of views, "instruct[ed] Jan. 6 rally attendees to send

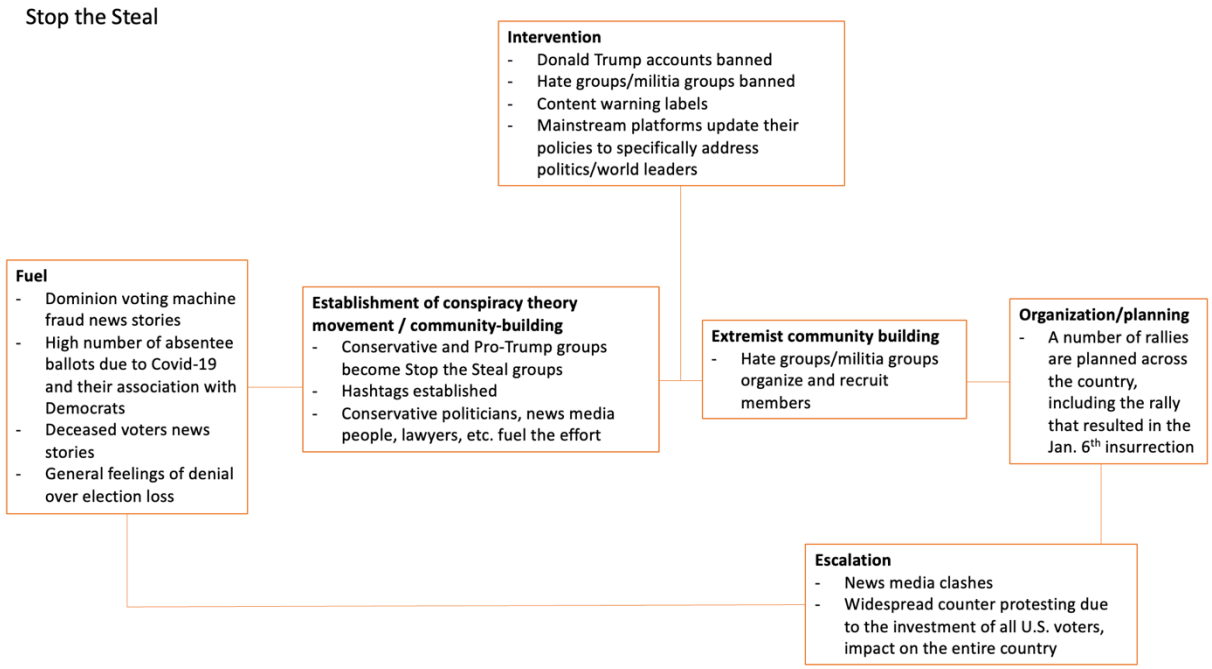
locations and pictures of “BLM and Antifa” to the Proud Boys so that members of the group can “get them before they go out to the streets,”” and “Posts became increasingly violent on Parler and other message boards in the days leading up to January 6” (DFRLab, 2021).

Foreign Intervention

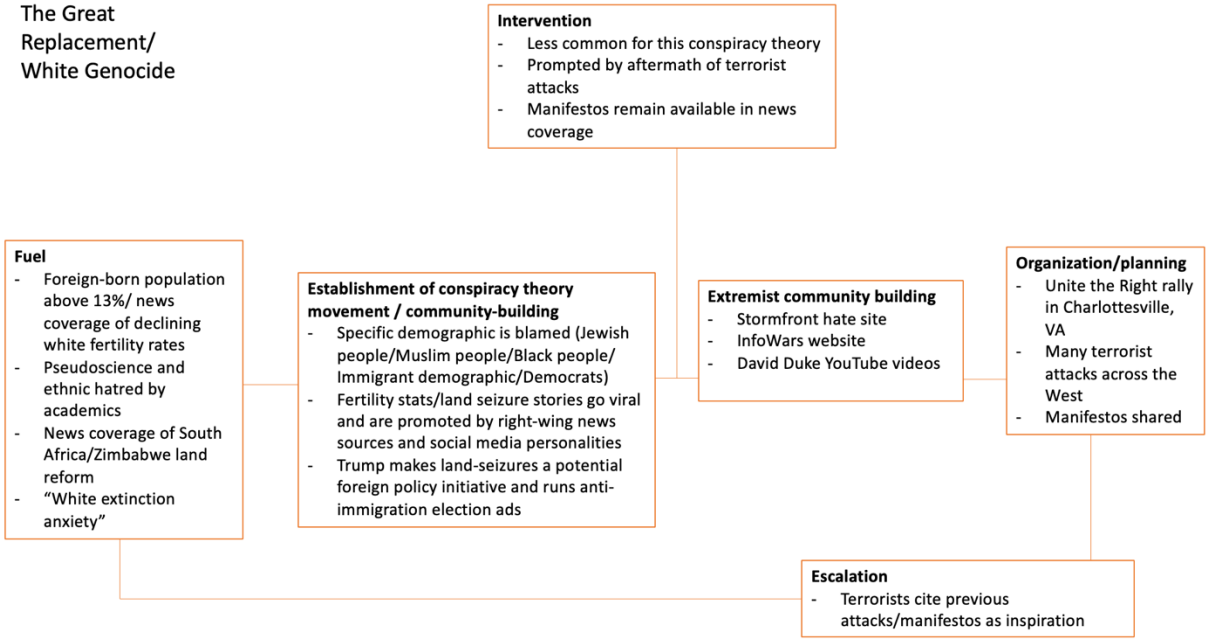
Political adversaries of the United States often seek to capitalize on internal political strife. An Iranian disinformation campaign attempted to capitalize on Stop the Steal by circulating an “enemies list” of politicians, cyber experts, and election officials who did not support election fraud claims.

Flowcharts

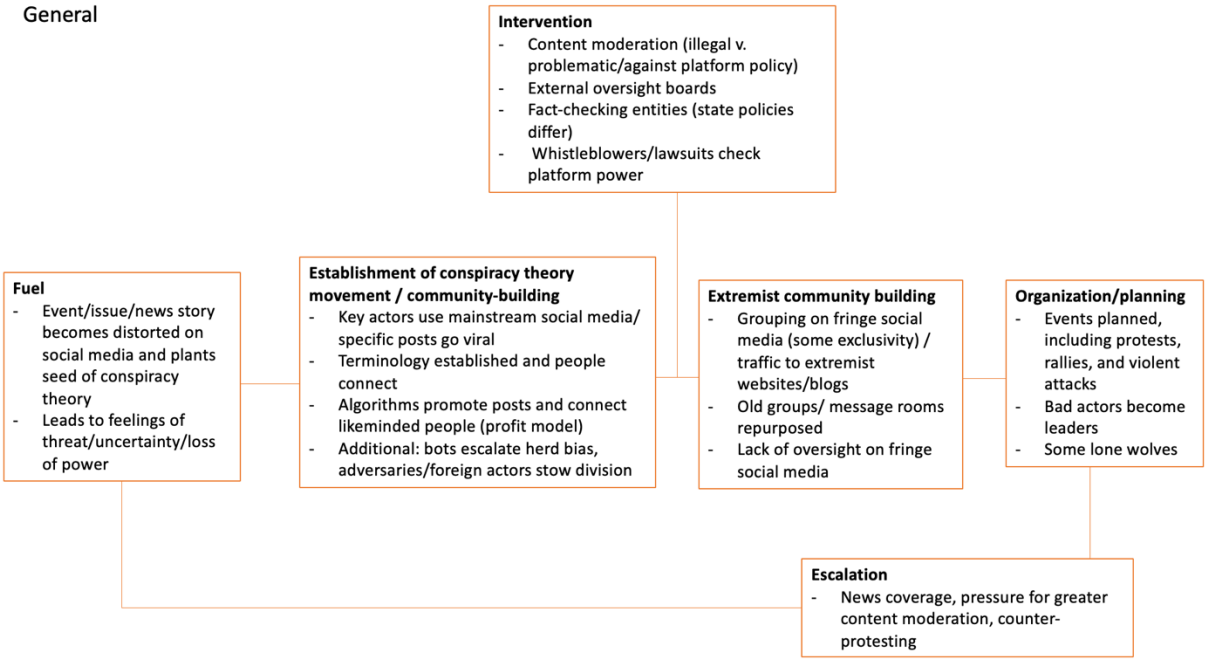
These flowcharts synthesize previous findings and provide a visual representation of the flow of conspiracy theories across social media platforms.



The Great Replacement/ White Genocide



General



Chapter 5: Conclusion

The Impact of Network Models on the Evolution of Conspiracy Theories

The findings in Chapter 4 indicate that the differing models and content moderation practices of mainstream and fringe platforms provide different advantages and disadvantages that are capitalized on by those intending to spread a conspiracy theory. Mainstream platforms lend the ability to reach a massive number of users in an instant, before even successful content moderation practices are able to flag or remove posts. Fringe platforms, on the other hand, have the advantage to conspiracy theorists of employing little to no moderation, particularly when the platforms themselves market as right-wing communities or "free-speech havens." This is why conspiracy theorists use mainstream platforms to reach many users and then direct them to fringe platforms or even personal blogs and websites. This leads conspiracy theories to flow from mainstream to fringe platforms, where extremist groups already reside in communities, and where the conspiracy theory is escalated.

Right-wing news media also makes use of mainstream platforms and often echoes the sentiment of right-wing conspiracy theories on social media. This harmonious interplay between news media dialogue and social media dialogue appears to be unique to conservative sentiment and provides a sense of legitimacy to the conspiracy theories. Even anonymous posters on social media have the chance and ability to be backed by massive news networks.

Policy Implications

The key actors findings indicate that platforms with a motivation to curb the spread of misinformation may find better success focusing content moderation efforts on identified key spreaders. These are people with a large audience- often politicians, right-wing news media actors, popular bloggers or hosts of podcasts and talk shows, and even celebrities who interact

heavily with this type of content. A similar focus can be directed toward groups that have been identified as spreading mis/disinformation or extremist content in the past, as these groups are often repurposed to support and spread new and related causes.

The flow from mainstream to fringe findings indicate that mainstream social media platforms should monitor the external websites, blogs, chatrooms, and outside platforms that are linked by prominent figures or by popular posts. Mainstream platforms could attempt to remove posts or block external links when these external sources promote or organize violence. Mainstream platforms are powerful organizing tools because of their sheer number of users, and conspiracy theorists take advantage of this tool while circumventing content moderation by then sending users to less-monitored sites or platforms. This flow of users could also be monitored by the mainstream platforms themselves by identifying heavily clicked links and moderating traffic if the outside sources are deemed to violate the guidelines of the original platform. This could take the form of removing posts, blocking links, or enacting content warning labels.

Generally, the findings show an attempt by mainstream social media platforms, particularly in the wake of violence related to Stop the Steal, to address disinformation and incitement of violence on their platforms and outline specific guidelines for this type of content. Fringe platforms, on the other hand, appear to completely lack or avoid oversight, even for unprotected speech such as incitement of violence. This indicates that there may not be a need for sweeping state or federal technology policy changes, but rather greater scrutiny of users that violate existing laws. However, in terms of legislative changes, ascribing a level of responsibility to platforms that allow incitement to violence to go unchecked would likely be successful in suppressing the spread of much of this content by motivating all platforms to more closely

moderate. As mentioned in Chapter 2, the primary option is to designate platforms as publishers of content rather than intermediaries, particularly when their algorithms amplify such content.

Suggestions for Future Research

Fringe platforms seem to be increasingly popular and tend to posit themselves as existing for a particular purpose, often having their own political affiliation. Their emergence, rise in popularity, and often subsequent fall, is a process that could be valuable to study further. Hate groups have found footholds on many of these platforms, where they continue to build their online community. The further examination of how specific extremist groups use social media to grow and amplify their message is an important initiative for both technology policy analysts and law enforcement entities.

References

Arango, T., Bogel-Burroughs, N., & Benner, K. (2019, August 4). *Minutes Before El Paso Killing, Hate-Filled Manifesto Appears Online - The New York Times*. The New York Times - Breaking News, US News, World News and Videos. https://www.nytimes.com/2019/08/03/us/patrick-crusius-el-paso-shooter-manifesto.html

Betuel, E. (2019, January 13). *CDC Data on Declining US Fertility Rate Is Being Used by White Nationalists*. Inverse; Inverse. https://www.inverse.com/article/52409-cdc-fertility-report-race-breakdown-white-nationalism

Board, T. E. (2021, August 11). U.S. Capitol combat shows why police deaths by suicides are sacrifices in the line of duty. Retrieved from https://www.usatoday.com/story/opinion/todaysdebate/2021/08/11/us-capitol-police-jan-6-officers-died-suicide-line-duty/5541978001/

Bond, S. (2021). *Trump Suspended From Facebook For 2 Years*. Npr. https://www.npr.org/2021/06/04/1003284948/trump-suspended-from-facebook-for-2-years

Bond, S., & Allyn, B. (2021). *How the "Stop the Steal" movement outwitted Facebook ahead of the Jan. 6 insurrection*. Npr. https://www.npr.org/2021/10/22/1048543513/facebook-groups-jan-6-insurrection

Breland, A. (2020a). *Trump's Latest False Fraud Claim Came From an Anonymous Internet Poster – Mother Jones*. Mother Jones. https://www.motherjones.com/politics/2020/11/trump-dominion-fraud-oann/

Breland, A. (2020b, October 21). *How the Most Toxic, Notorious Pro-Trump Online Community Tricked Reddit and Got Back Online – Mother Jones*. Mother Jones. https://www.motherjones.com/politics/2020/10/the_donald_reddit_discord/

Bullens, L. (2021, November 08). How France's 'great replacement' theory conquered the global far right. Retrieved from https://www.france24.com/en/europe/20211108-how-the-french-great-replacement-theory-conquered-the-far-right

Coaston, J. (2018, January 18). *The white genocide conspiracy theory, explained - Vox*. Vox; Vox. https://www.vox.com/2018/1/18/16897358/racism-donald-trump-immigration

Congressional Research Service. (2021, January 27). Social Media: Misinformation and Content Moderation Issues for Congress. https://crsreports.congress.gov/product/pdf/R/R46662/6

Daly, K. (2020, July 27). 4 numbers that show the combined power of Facebook, Google, Apple and Amazon. Retrieved from https://www.axios.com/big-techs-power-in-4-numbers-de8a5bc3-65b6-4064-a7cb-3466c68b2ea0.html

Dean, B. (2021, October 10). How Many People Use Social Media in 2021? (65 Statistics). Retrieved from https://backlinko.com/social-media-users

Demographics of Social Media Users and Adoption in the United States. (2021, November 23). Retrieved from https://www.pewresearch.org/internet/fact-sheet/social-media/

DFRLab. (2021, February 10). *#StopTheSteal: Timeline of Social Media and Extremist Activities Leading to 1/6 Insurrection*. Just Security; Atlantic Council. https://www.justsecurity.org/74622/stopthesteal-timeline-of-social-media-and-extremist-activities-leading-to-1-6-insurrection/

Draper, N. (2019) Distributed intervention: networked content moderation in anonymous mobile spaces. *Feminist Media Studies*, 19:5, 667-683, DOI: 10.1080/14680777.2018.1458746

Easterbrook, T. (2021). *The Rhetoric of Conspiracism in User-Centered Democracy.* https://doi.org/10.17615/qtv1-7489

Emerson, S. (2020a, November 20). *'Stop the Steal' Groups Are Still Thriving on Facebook | OneZero*. Medium; OneZero. https://onezero.medium.com/stop-the-steal-groups-are-still-thriving-on-facebook-330a388a2901

Emerson, S. (2020b, November 10). *Trump Supporters Flock to Parler, MeWe, and Rumble Following Facebook Ban | OneZero*. Medium; OneZero. https://onezero.medium.com/trump-supporters-flock-to-free-speech-platforms-after-facebook-ban-bf3e6ab1a70d

"Fake News," Lies and Propaganda: How to Sort Fact from Fiction. (2021). *University of Michigan Library Research Guides.* Retrieved from https://guides.lib.umich.edu/fakenews

Gillespie, T. (2018). Platforms Are Not Intermediaries. *Georgetown Law Technology Review, 2 Geo. L. Tech. Rev. 198.* https://georgetownlawtechreview.org/wp-content/uploads/2018/07/2.2-Gilespie-pp-198-216.pdf

Healy, J. (2021, January 11). These Are the 5 People Who Died in the Capitol Riot. Retrieved from https://www.nytimes.com/2021/01/11/us/who-died-in-capitol-building-attack.html

Klintman, M. (2019, December 13). *Conspiracy theories: how belief is rooted in evolution – not ignorance*. The Conversation; https://facebook.com/ConversationUK. https://theconversation.com/conspiracy-theories-how-belief-is-rooted-in-evolution-not-ignorance-128803

Lajka, A. (2020, November 10). *Election breathes new life into false "dead voter" claims*. AP News. https://apnews.com/article/ap-fact-check-election-dead-voter-claims-fc7ba254fd37059f63ea764c18c2a4cb

LSE Commission on Truth, Trust, and Technology. (2019, July 30). Tackling the Information Crisis: A Policy Framework for Media System Resilience. *The London School of Economics and Political Science.* https://perma.cc/5Z7S-FJU7.

Lytvynenko, J., & Hensley-Clancy, M. (2021, January 6). *Rioters Taking Over The Capitol Planned It Online*. BuzzFeed News; BuzzFeed News. https://www.buzzfeednews.com/article/janelytvynenko/trump-rioters-planned-online?scrolla=5eb6d68b7fedc32c19ef33b4

MacMillan, D. (2022, April 27). *Elon Musk's 'free speech' takeover part of New Corporate Activism Wave*. The Washington Post. Retrieved April 27, 2022, from https://www.washingtonpost.com/business/2022/04/27/twitter-free-speech-elon-musk/

Marwick, Alice E. "Why Do People Share Fake News? A Sociotechnical Model of Media Effects." *Georgetown Law Technology Review*, 478, 2018, https://georgetownlawtechreview.org/wp-content/uploads/2018/07/2.2-Marwick-pp-474-512.pdf.

Nunberg, G. (2019, December 30). 'Disinformation' Is The Word Of The Year - And A Sign Of What's To Come. Retrieved from https://www.npr.org/2019/12/30/790144099/disinformation-is-the-word-of-the-year-and-a-sign-of-what-s-to-come

Permanent suspension of @realDonaldTrump. (2021, January 8). Blog. Retrieved April 6, 2022, from https://blog.twitter.com/en_us/topics/company/2020/suspension

Pogue, J. (2019, February 15). *The Myth of White Genocide | Pulitzer Center*. Pulitzer Center. https://pulitzercenter.org/stories/myth-white-genocide

Reps. Malinowski and Eshoo Reintroduce Bill to Hold Tech Platforms Accountable for Algorithmic Promotion of Extremism. (2021, March 24). Retrieved from https://malinowski.house.gov/media/press-releases/reps-malinowski-and-eshoo-reintroduce-bill-hold-tech-platforms-accountable

Rivero, N. (2020, November 2). The risk of putting warning labels on election misinformation. *Quartz.* https://qz.com/1925272/the-risk-of-putting-warning-labels-on-election-misinformation/

Russo, A. (2019, August 6). *Trump Ran Facebook Ads Decrying An Immigrant "Invasion" | HuffPost Latest News*. HuffPost; HuffPost. https://www.huffpost.com/entry/donald-trump-facebook-ads-immigration_n_5d49b208e4b0d291ed07a5af

Seetharaman, D. (2020, Dec 09). Facebook antitrust lawsuits: What do the cases mean for the social-media giant? U.S. antitrust authorities have been examining for more than a year whether Facebook used its market position to stifle competition. *Wall Street Journal (Online)* Retrieved from http://libproxy.lib.unc.edu/login?url=https://www-proquest-

com.libproxy.lib.unc.edu/newspapers/facebook-antitrust-lawsuits-what-do-cases-mean/docview/2468411750/se-2?accountid=14244

Shahsavari, S., Holur, P., Wang, T. *et al.* Conspiracy in the time of corona: automatic detection of emerging COVID-19 conspiracy theories in social media and the news. *J Comput Soc Sc* 3, 279–317 (2020). https://doi.org/10.1007/s42001-020-00086-5

Spencer, S. H. (2020, November 25). *Baseless Conspiracy Theory Targets Another Election Technology Company - FactCheck.org*. FactCheck.Org; https://www.facebook.com/factcheck.org. https://www.factcheck.org/2020/11/baseless-conspiracy-theory-targets-another-election-technology-company/

Stubbersfield, J. M. (2021, September 2). Conspiracy Theories: A Cultural Evolution Theory approach. https://doi.org/10.31219/osf.io/hqkju

Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary. (2020). Competition in Digital Markets - United States House of ... Retrieved from https://judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf

The Associated Press. (2018, October 31). *Far-right group Proud Boys banned from Facebook, Instagram | CBC News*. CBC. https://www.cbc.ca/news/canada/proud-boys-facebook-instagram-1.4886514

United States General Accounting Office, Program Evaluation and Methodology Division (1987). Case Study Evaluations. Retrieved from https://www.gao.gov/assets/132683.pdf

van Gerven Oei, V. W. J. (2011). *Anders Breivik: On Copying the Obscure | continent.* Wayback Machine. https://web.archive.org/web/20200716125213/http://www.continentcontinent.cc/index.php/continent/article/view/56

van Mulukom, V. (2021, September 4). Dual pathway model of conspiracy beliefs: Interactions between epistemic and socio-existential needs and resources. https://doi.org/10.31234/osf.io/z5b4f

van Prooijen, J.-W. (2019). *An Existential Threat Model of Conspiracy Theories*. Hogrefe Publishing Group. https://econtent.hogrefe.com/doi/epdf/10.1027/1016-9040/a000381

Waterson, J. (2022a, January 4). *Twitter permanently bans news aggregation service Politics For All | Twitter | The Guardian*. The Guardian; The Guardian. https://www.theguardian.com/technology/2022/jan/04/twitter-permanently-bans-news-aggregation-service-politics-for-all

World Leaders on Twitter: principles & approach. (2019, October 15). Blog. Retrieved April 6, 2022, from https://blog.twitter.com/en_us/topics/company/2019/worldleaders2019